



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
-----------------	-------------	----------------------	---------------------	------------------

10/729,889

12/05/2003

Todd D. Wakefield

5133

9228

64100

7590

10/31/2007

DANIEL P. MCCARTHY

P.O. BOX 71550

SALT LAKE CITY, UT 84171-0550

EXAMINER

TIMBLIN, ROBERT M

ART UNIT

PAPER NUMBER

2167

MAIL DATE

DELIVERY MODE

10/31/2007

PAPER

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

41

<b>Office Action Summary</b>	<b>Application No.</b> 10/729,889	<b>Applicant(s)</b> WAKEFIELD ET AL.	
	<b>Examiner</b> Robert M. Timblin	<b>Art Unit</b> 2167	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☒ Responsive to communication(s) filed on 23 August 2007.
- 2a) ☐ This action is **FINAL**.                      2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 1,4-18 and 21-34 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1,4-18 and 21-34 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- |   |   |
|---|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892)  | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)  | 5) <input type="checkbox"/> Notice of Informal Patent Application                       |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)<br>Paper No(s)/Mail Date <u>10/23/2007</u> | 6) <input type="checkbox"/> Other: _____  |

### **DETAILED ACTION**

This office action corresponds to application 10/729,889 filed 12/05/2003.

Response to arguments begin on page 15 of this Office Action.

#### ***Continued Examination Under 37 CFR 1.114***

A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 8/23/2007 has been entered.

#### ***Terminal Disclaimer***

The terminal disclaimer filed on 11/19/2006 is accepted with the corrections made thereto.

#### ***Response to Amendment***

In accordance with the present amendments, claims 1, and 17-18 are amended. Accordingly, claims 1, 4-18 and 21-34 are pending.

The previous claim objections have been removed.

### ***Double Patenting***

The nonstatutory double patenting rejection is based on a judicially created doctrine grounded in public policy (a policy reflected in the statute) so as to prevent the unjustified or improper timewise extension of the "right to exclude" granted by a patent and to prevent possible harassment by multiple assignees. A nonstatutory obviousness-type double patenting rejection is appropriate where the conflicting claims are not identical, but at least one examined application claim is not patentably distinct from the reference claim(s) because the examined application claim is either anticipated by, or would have been obvious over, the reference claim(s). See, e.g., *In re Berg*, 140 F.3d 1428, 46 USPQ2d 1226 (Fed. Cir. 1998); *In re Goodman*, 11 F.3d 1046, 29 USPQ2d 2010 (Fed. Cir. 1993); *In re Longi*, 759 F.2d 887, 225 USPQ 645 (Fed. Cir. 1985); *In re Van Ornum*, 686 F.2d 937, 214 USPQ 761 (CCPA 1982); *In re Vogel*, 422 F.2d 438, 164 USPQ 619 (CCPA 1970); and *In re Thorington*, 418 F.2d 528, 163 USPQ 644 (CCPA 1969).

A timely filed terminal disclaimer in compliance with 37 CFR 1.321(c) or 1.321(d) may be used to overcome an actual or provisional rejection based on a nonstatutory double patenting ground provided the conflicting application or patent either is shown to be commonly owned with this application, or claims an invention made as a result of activities undertaken within the scope of a joint research agreement.

Effective January 1, 1994, a registered attorney or agent of record may sign a terminal disclaimer. A terminal disclaimer signed by the assignee must fully comply with 37 CFR 3.73(b).

Claims 1, 4-18 and 21-34 are provisionally rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claims 1, 4-6, 9, 10, 14-15, and 30-31 of copending Application No. 10/729,864. This is a provisional obviousness-type double patenting rejection because the conflicting claims have not in fact been patented.

Claims 1, 4-18 and 21-34 are provisionally rejected on the ground of nonstatutory obviousness-type double patenting as being unpatentable over claims 1, 2, 5-17, 20-33, 36-47 of copending Application No. 10/728721. This is a provisional obviousness-type

Art Unit: 2167

double patenting rejection because the conflicting claims have not in fact been patented.

### ***Claim Objections***

Claims 4-16 and 21-34 are objected to under 37 CFR 1.75(c), as being of improper dependent form for failing to further limit the subject matter of a previous claim. Applicant is required to cancel the claim(s), or amend the claim(s) to place the claim(s) in proper dependent form, or rewrite the claim(s) in independent form.

These claims should start with the word "The" rather than "A" so that the depending claims clearly incorporate and further limit the previous claim.

### ***Claim Rejections - 35 USC § 102***

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

Claims 17, 18, and 21-27 are rejected under 35 U.S.C. 102(b) as being anticipated by **Gaizauskas et al.** "Information Extraction: Beyond Document Retrieval" August 1998. ('**Gaizauskas**' hereinafter). In the following passages and figures, **Gaizauskas** teaches:

Art Unit: 2167

With respect to claim 17, a computer program product located to one or more storage media devices usable to perform integration of mixed format data, said computer program product comprising instructions executable by a computer to perform the functions of:

accessing a database containing data records (page 18; retrieving documents from collections, page 48; newsfeeds, page 45 last paragraph; running text), at least some of the data records containing both (figure 1b; e.g. the tagged elements between <DOC> and <SO> are structured data) data and unstructured data (figure 1, step b; i.e. the free text included and tagged by <p>), the unstructured data including free text (introduction, page 17) that is added to the data records independently of the structured data (i.e. free text is seen as independent from structured data), wherein the structured data reflects information captured separately from the unstructured data (page 18; retrieving documents from collections. Also, free text suggests entering data separately from structured data),

using linguistic information to extract relational facts from the free text (beginning of page 19, first full paragraph of page 47; linguistic analysis/theory);

producing a set of construed data (figure 1d) reflecting at least one relational fact (page 42, 6<sup>th</sup> line of the first full paragraph) conveyed in said free text (i.e. text between <p> and </p>, each construed datum in the set of construed data containing at least one relational fact;

relating each construed datum to the structured data in the corresponding data record in which the corresponding free text was stored (i.e. figure 1d describes relating the construed data to the <DOCNO> of the structured data);

Art Unit: 2167

integrating said construed data with the structured data of the data record to which said construed data relates, said integrating step retaining reference information to the original free text (page 39 section 3.2.3 and second full paragraph of page 44); and

constructing a library containing extracted attributes (figures 2 and 4).

With respect to claim 18, a method for integrating mixed format data, comprising the steps of:

accessing a database containing data records, at least some of the data records containing both structured and unstructured data, the unstructured data including free text (page 18; retrieving documents from collections) that is generated by first processes that are separate from second processes that generate the structured data (i.e. free text, as disclosed by Gaizauskas, suggest data that is entered as a separate process that generates the structured data);

producing a set of construed data (figure 1d) reflecting at least one relational fact (e.g. organizational information, in\_and\_out information and succession event information) conveyed in free text (text in figure 1 tagged by <p>), each construed datum (figure 1d) containing at least one relational fact (e.g. organizational information, in\_and\_out information and succession event information);

relating each construed datum with the data records in which corresponding free text is located (i.e. figure 1d relates the construed data to <DOCNO> of the original structured data); and

integrating the produced data the structured data (page 39 section 3.2.3 and second full paragraph of page 44).

With respect to claim 21, a method according to claim 18, further comprising the step of applying caseframes to said free text (last paragraph of page 22, first paragraph of page 23, and figure 3 on page 39).

With respect to claim 22, a method according to claim 18, further comprising the step of producing a new database containing the integrated data produced by said integrating step (page 1, and page 50 section 5.1.3).

With respect to claim 23, a method according to claim 18, further comprising the step of inserting the produced data into said database (introduction, first paragraph).

With respect to claim 24, a method according to claim 18, further comprising the step of creating a new database (figure 1 on page 20 and section 5.1.3.).

With respect to claim 25, a method according to claim 24, wherein the new database is a relational database (page 52; conventional database and section 5.1.3.).

With respect to claim 26, a method according to claim 24, wherein new database includes at least one file containing the integrated data produced by said integrating step (number 2 on page 29 and figure 2 on page 35).



With respect to claim 27, a method according to claim 26, wherein the new database has a format selected from the group of XML, character separated values, spreadsheet formats and file-based database structures (figure 1 on page 20 and number 2 on page 29).

With respect to claim 28, a method according to claim 18, further comprising the step of combining like attributes for the extracted relational facts produced in performing said extracting relational facts from the free text (figures 4-5 and accompanying descriptions).

With respect to claim 29, a method according to claim 18, further comprising the step of combining like relation types for the extracted relational facts produced in performing said extracting relational facts from the free text (introduction, first paragraph).

With respect to claim 30, a method according to claim 18, wherein domain roles are applied in said step of extracting relational facts from the free text (introduction, first paragraph).

With respect to claim 31, a method according to claim 18, further comprising the step of storing the relational facts produced in performing said extracting relational facts from the free text (page 22, last paragraph).

With respect to claim 32, a method according to claim 18, wherein the extracted relational facts produced in performing said extracting relational facts and the integrated data produced by the performance of said integrating the produced data includes reference information to the original free text (figure 3 and accompanying description).

With respect to claim 33, a computer program product according to claim 1, wherein said instructions are further executable to replace like or related attributes for relational facts with a common canonical representation based on those like or related attributes (first paragraph of page 19, first paragraph of page 27, and page 30).

With respect to claim 34, a computer program product according to claim 1, wherein said instructions are further executable to replace like or related relation fact types with a common canonical representation.

### ***Claim Rejections - 35 USC § 103***

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Claims 1 and 4-16 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gaizauskas as applied to claims 17, 18, and 21-27 above in view of Alpha et al. ("Alpha" hereafter) U.S. Patent 6,980,976 B2.

Art Unit: 2167

With respect to claim 1, a computer program product located to one or more storage media devices usable to perform integration of mixed format data, said computer program product comprising instructions executable by a computer to perform the functions of:

accessing a feed of database records (page 18; retrieving documents from collections, page 48; newsfeeds, page 45 last paragraph; running text), said database records including both structured (figure 1b; e.g. the tagged elements between <DOC> and <SO> are structured data) data and unstructured data (figure 1, step b; i.e. the free text included and tagged by <p>);

the unstructured data (free text) of a particular database record ((figure 1) including free text (i.e. figure 1b discloses retrieved text from an article and furthermore, the introduction teaches populating from a free text source)), wherein the unstructured data (free text) is derived independently of the structured data (i.e. in Gaizauskas' introduction, it is disclosed that their teachings are not exclusive to articles, and may include free text as an information source. "Free text" is therefore seen as independent from structured data);

extracting relational facts (figure 1c/d; i.e. Gaizauskas discloses extracting the free text and placing key elements in a template. The template includes relational information (e.g. organizational information, in\_and\_out information and succession event information)) from the free text (page 27, paragraph 3 and page 29), said extracting step being performed using linguistic information from the free text (beginning of page 19, figure 3, first full paragraph of page 47; linguistic analysis/theory);

producing a set of construed data (figure 1 d; i.e. the data from the free text is organized according to the template. Also see page 30, last paragraph where Gaizauskas teaches an example of semantic interpretation to construe data) from said unstructured data (figure 1b; free text), each construed datum in the set of construed data containing at least one relational fact (step d) of figure 1, at least steps 4-7 on page 34, and processing stages on page 36 and figure 3 on page 39 with description),

relating each construed datum (for example, page 42, line 6 of the first full paragraph shows a fact) to the structured data of the database record in which said free text was found (steps b and d of figure 1). Furthermore, Gaizauskas teaches in figure 1 that the free text is extracted and placed into a template (figure 1d) which includes relating each construed datum to the <DOCNO> of the original structured data); and

integrating the construed data (figure 1d) with the particular structured data to which the construed data relates (page 39 section 3.2.3 and second full paragraph of page 44).

Gaizauskas, fails to expressly teach a feed of relational database records.

Alpha, however, teaches a feed relational database records (drawing reference 105 and figure 3) to provide a data source including structured data columns and unstructured data columns.

In the same field of endeavor, (i.e. information extraction), it would have been obvious to one of ordinary skill in the data processing art at the time of the present invention to combine the teachings of the cited references because Alpha providing Gaizauskas with relational database records would have given their teachings

Art Unit: 2167

structured input (i.e. columns) for efficient processing and further a wide range of input data for processing.

With respect to claim 4, a computer program product according to claim 1, further comprising the step of applying caseframes while performing said extracting step (last paragraph of page 22, first paragraph of page 23, and figure 3 on page 39).

With respect to claim 5, a computer program product according to claim 1, wherein said instructions are further executable to perform the function of producing a new database containing the integrated data produced by said integrating (page 1, and page 50 section 5.1.3).

With respect to claim 6, a computer program product according to claim 1, wherein said data feed is a database, and wherein said instructions are further executable to perform the function of inserting the construed data into said database while performing said integrating step (introduction, first paragraph).

With respect to claim 7, a computer program product according to claim 1, wherein said instructions are further executable to perform the function of creating a new database containing the construed data (figure 1 on page 20 and section 5.1.3.).

With respect to claim 8, a computer program product according to claim 7, wherein said new database is a relational database which relates said relational facts to said structured data (page 52; conventional database and section 5.1.3.).

With respect to claim 9, a computer program product according to claim 8, wherein the instructions are further executable to produce a file containing the integrated data produced by said integrating (number 2 on page 29 and figure 2 on page 35).

With respect to claim 10, a computer program product according to claim 9, wherein the instructions are further executable to produce a file having a format selected from the group of XML, character separated values, spreadsheet formats and file-based database structures (figure 1 on page 20 and number 2 on page 29).

With respect to claim 11, a computer system including a computer program product according to claim 1, further comprising: a processing unit coupled to said one or more storage media devices, said processing unit being capable of executing said instructions; and an execution command unit, whereby operation of said instructions and said processing unit may be commanded or controlled (page 46, first full paragraph).

With respect to claim 12, a computer program product according to claim 1, wherein said instructions are further executable to combine like attributes for the

Art Unit: 2167

extracted relational fact types produced in performing said extracting relational facts from the free text (figures 4-5 and accompanying descriptions).

With respect to claim 13, a computer program product according to claim 1, wherein said instructions are further executable to combine like relational fact types for the extracted relational facts produced in performing said extracting relational facts from the free text (first paragraph of 5.1.2).

With respect to claim 14, a computer program product according to claim 1, wherein said instructions provide relationships with domain roles applied in performing said extracting relational facts from the free text (page 22, last paragraph).

With respect to claim 15, a computer program product according to claim 1, wherein said instructions store the relational facts produced in performing said extracting relational facts from the free text (introduction, first paragraph).

With respect to claim 16, a computer program product according to claim 1, wherein the extracted relational facts produced in performing said extracting relational facts and the integrated data produced by the performance of said integrating the produced data includes reference information to the original free text (figure 3 and accompanying description).

***Response to Arguments***

Applicant's arguments filed 8/23/2007 have been fully considered but they are not persuasive.

The Applicant argues on page 10-13 of the remarks that Gaizauskas does not teach wherein the unstructured data is derived independently of the structured data. The Examiner respectfully disagrees because as seen in figure 1b, Gaizauskas discloses a data record including structured data (i.e. the tagged elements between <DOC> and <SO> are structured data) and unstructured data (i.e. the text marked up by <p> </p>). That is, the text found within the structured portion is natural, free language text (i.e. unstructured data). Gaizauskas explicitly teaches populating a structured information source from an unstructured information source (introduction, lines 7-8). Gaizauskas further teaches the unstructured source includes but is not restricted to newswire texts and may include free text (introduction, lines 3-8).

Aside from Gaizauskas teaching the unstructured data as "free text" to suggest it's independence from structured data (i.e. "free" text is seen to be independent from structured data), Gaizauskas also suggests the unstructured data's independence by the populating of a structured source. The Applicant states that one way to recognize the distinction [of structured data derived independently] is to identify when the structured data are created and assigned to the unstructured content (Applicant's remarks, last paragraph of page 11). The Examiner respectfully submits that when Gaizauskas teaches populating a structured information source with unstructured data,



• Art Unit: 2167

the structures would already have been defined (i.e. independent from unstructured free text) in order for population of free text into the structures to take place.

The Applicant further argues (page 13, 3<sup>rd</sup> paragraph) that Gaizauskas does not teach relating the construed datum back to the independent structured data of a database record from which the free text was drawn as required by claim 1.

The Examiner respectfully disagrees because Gaizauskas teaches retrieving text (figure 1b) and subsequently filling a template (figure 1c-d). As can be seen in the template (figure 1d), data from the unstructured passage retrieved (figure 1b), parsed, and interpreted (i.e. construed) into template form. The template (1d) shows that data from the passage are extracted to produce construed data. For example, "Donald Wright" is extracted and construed as being a person. This information is placed accordingly as a person in the template as construed datum. Furthermore, it also be seen that each construed datum is related back to the original record by way of the document number 940413-0062. Specifically, each set of construed data (i.e. succession\_event, in\_and\_out, organization, and person) is related back to the record in which the free text was found by containing that document number.

For at least these reasons, the present invention is not patentable over the prior art.

Art Unit: 2167

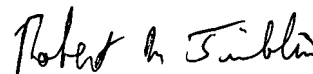
### Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Robert M. Timblin whose telephone number is 571-272-5627. The examiner can normally be reached on M-F 8:00-4:30.

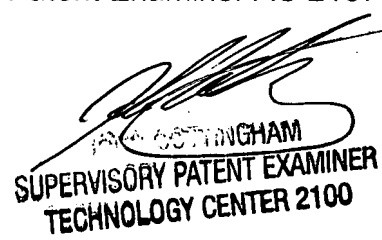
If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John R. Cottingham can be reached on 571-272-7079. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Robert M. Timblin



Patent Examiner AU 2167



JOHN R. COTTINGHAM  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 2100